Design and Analysis of a Hash Ring-iterative Structure*

Shenghui Su[1], Yixian Yang[2], Bo Yang[2], and Shaolan Zhang[2]

[1]College of Computer Science, Beijing University of Technology, Beijing 100022, P. R. China

[2]School of Information Engineering, Beijing University of Posts & Telecom, Beijing 100876, P. R. China

Abstract: The authors propose a new type of hash iterative structure ─ the ring-iterative structure with feedback which is subdivided into the single feedback ring iteration and the multiple feedback ring iteration, namely SFRI and MFRI. Prove that SFRI is at least equivalent to the MD structure in security, and MFRI is at least equivalent to SFRI in security (property 1 makes people incline to believe MFRI is more secure than MD). Analyze the resistance of MFRI, which results from the joint event on message modification, endless loop on message modification and incompatibility of the sufficient conditions, to the multi-block differential collision attack. Argue the ineffectiveness of the D-way second preimage attack on MFRI. Discuss the time and space expenses of MFRI, and point out the advantage of MFRI over the tree-iterative structure and the zipper-iterative structure.

Keywords: Digital signature, Hash function, Security, Ring Iteration, Compression function.

1 Introduction

It is well known that hash functions are primarily employed for digital signature, data integrity and message authentication code which are widely used in trust computing systems. The security of hash functions is the foundation of security of digital signature.

At present, almost all famous hash functions ─ MD5, SHA-0, and SHA-1 [1] for example adopt the Merkle-Damgård (MD) iterative structure [2][3]. The design principle of this structure is that if there does not exist a computationally collision-free function *h* mapping a message of arbitrary polynomial length to a *k*-bit string, then there does not exist a computationally collision-free function *f* from *m* bits to *k* bits, where $k < m$ [2]. Thereby, it has been universally thought that the problem of designing a collision-free hash function may be reduced to the problem of designing a collision-free compression function, namely iterative function.

However, the multi-block differential collision attack on MD5、SHA-0, and SHA-1 [4][5][6] indicates that a collision-free compression function is not a sufficient condition of a collision-free hash function, but only a necessary condition [7]. It means that a secure and collision-free hash function will be based

* This research is supported by SCRL.

It occurs in *Chinese Journal of Electronics*, vol.20(2), Apr. 2010, pp. 232-236.

Corresponding email: reesse@126.com. Last revised Jan. 16, 2010.

not only on a collision-free compression function, but also on a collision-free iteration structure.

2 Design of Hash Ring-iterative Structures

2.1 Single Feedback Ring Iteration

Assume that a message to be hashed is X of l-bit length, and X is partitioned into n m-bit blocks X_1, …, X_n, where $n = l / m$ and l is exactly divided by m, that is, the padding problem is neglected by us, which does not influence our discussion.

Let IV be the initial value of the chaining variable, f be a compression function, every iterative output be Y_i of k-bit, where $k \le m$, $i = 1, \ldots, n$, and D be the last output, namely the message digest.

For MD, there are $Y_0 = IV$, $Y_i = f(Y_{i-1}, X_i)$, and $D = Y_n$.

The single feedback ring iteration, shortly SFRI, is a simple structure. It feeds back the reverse of the MD output Y_n into iterative box 1, sends the second output of iterative box 1 to iterative box n, and generates the message digest D last. See Figure 1.

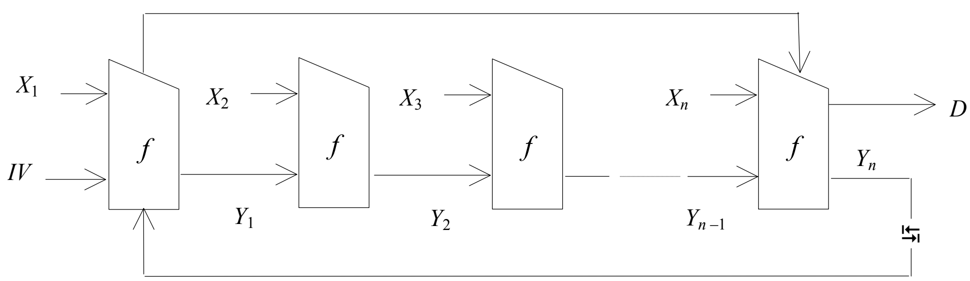

Figure 1: The Single Feedback Ring-iterative Structure

In Figure 1, we define $X_{n+1} = X_1$, $X_{n+2} = X_n$, and the message digest $D = f(X_{n+2}, f(X_{n+1}, ⇆Y_n))$.

Although the inputs X_1 and X_n of box 1 and box n are employed twice, and they have respectively two outputs, it is not incompatible in logicality according to the above definitions.

Note that sign '⇆' denotes reversal operation, that is, the bits of a variable are rearranged in reverse order. For example, the reverse of '100110' is '011001'.

2.2 Multiple Feedback Ring Iteration

The multiple feedback ring iteration, shortly MFRI, is comparatively complicated. It feeds back the modular sum of reverses of all iterative outputs into iterative box 1, sends the second output of iterative box 1 to iterative box n, and generates the message digest D last. See Figure 2.

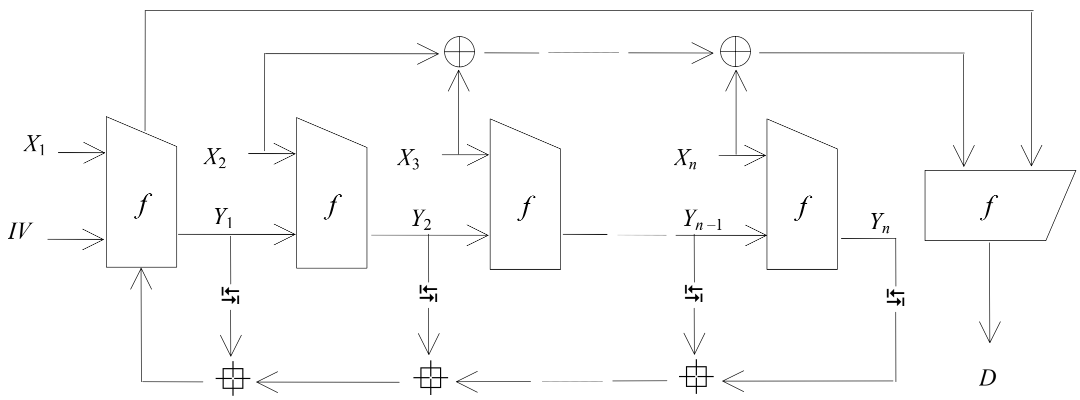

Figure 2: The Multiple Feedback Ring-iterative Structure

According to the above diagram, we define $X_{n+1} = X_1$, $X_{n+2} = X_2 \oplus \ldots \oplus X_n$, and the message digest $D = f(X_{n+2}, f(X_{n+1}, \leftrightarrows Y_1 \boxplus \ldots \boxplus \leftrightarrows Y_n))$, where sign '$\boxplus$' denotes modular addition operation.

Here, we substitute $X_{n+2} = X_2 \oplus \ldots \oplus X_n$ for $X_{n+2} = X_n$ in SFRI to make the last D depend relatively uniformly on $X_1, \ldots, X_n$. $X_2 \oplus \ldots \oplus X_n$ may be regarded as a feedforward.

3 Security Analysis of Ring-iterative Structures

3.1 At Least Equivalent to MD Structure in Security

Assume that f is a compression function. It combines with any iteration structure to construct a one-way hash function H.

***Definition* 1:** For a given message M, if we can not find in polynomial time a message $M' \neq M$ such that $H(M') = H(M)$, then H is called weakly collision-free.

***Definition* 2:** If we can not find in polynomial time any two messages M and M' satisfying $M' \neq M$ and $H(M') = H(M)$, then H is called strongly collision-free.

Obviously, if a hash function is strongly collision-free, then it must be also weakly collision-free.

3.1.1 SFRI Structure Being at Least Equivalent to MD Structure in Security

We temporarily neglect the existence of the operator $\leftrightarrows$ in SFRI. The SFRI without the operator $\leftrightarrows$ is called the reduced one.

In cryptology, security is measured with time complexity of attack tasks, and thus, if either of two iteration structures is strongly collision-free, they both are said to equivalent in security.

***Theorem* 1:** The reduced SFRI is equivalent to MD in security.

Proof:

Let H_1 be a hash function constructed with f and MD, and H_2 be another hash function constructed with f and the reduced SFRI structure.

1) Hypothesize that H_1 is strongly collision-free.

We need to prove that H_2 is also strongly collision-free. Proof by contradiction.

Presume that H_2 is not strongly collision-free, namely we can find the two messages M and M' satisfying $M \neq M'$ and $H_2(M) = H_2(M')$ in polynomial time.

Suppose that M is exactly partitioned into n m-bit blocks $X_1, \ldots, X_n$, and M' is exactly partitioned into n' m-bit blocks $X'_1, \ldots, X'_{n'}$.

In terms of the reduced SFRI, there are

$$H_2(M) = f(X_n, f(X_1, Y_n)), \text{ and } H_2(M') = f(X'_{n'}, f(X'_1, Y'_{n'})),$$

where $Y_n = H_1(M)$, and $Y_{n'} = H_1(M')$.

Let $M_1 = M \parallel X_1 \parallel X_n$, and $M'_1 = M' \parallel X'_1 \parallel X'_{n'}$, where '$\parallel$' represents the concatenation of strings. By comparing the structures of H_1 and H_2, there are

$$H_2(M) = H_1(M_1), \text{ and } H_2(M') = H_1(M'_1).$$

Therefore, in polynomial time, we can find the two messages M_1 and M'_1 which satisfy $M_1 \neq M'_1$ and $H_1(M_1) = H_1(M'_1)$. It is in direct contradiction to the hypothesis, so H_2 is also strongly collision-free, which indicates that the security of H_2 is not less than the security of H_1.

2) Hypothesize that H_2 is strongly collision-free.

We need to infer that H_1 is also strongly collision-free.

According to [5] and [6], define the message block differential as $\Delta X_i = X'_i - X_i$, and the iterative output differential as $\Delta Y_i = Y'_i - Y_i$.

For the reduced SFRI, the collision differential characteristics will be $\Delta X_1, \ldots, \Delta X_n, \Delta X_{n+1}, \Delta X_{n+2}$, and $0 = \Delta Y_0, \Delta Y_1, \ldots, \Delta Y_n, \Delta Y_{n+1}, \Delta Y_{n+2} = 0$.

Note that $\Delta X_{n+1} = X'_1 - X_1$ and $\Delta X_{n+2} = X'_n - X_n$. Thus, $\Delta X_1 = \Delta X_{n+1}$ and $\Delta X_n = \Delta X_{n+2}$.

If attackers set $\Delta X_1 = \Delta X_n = \Delta Y_n = 0$, attack on the reduced SFRI degenerates to attack on MD.

Thereby, if H_1 is not strongly collision-free, H_2 is also not strongly collision- resistant. It is in direct contradiction to the hypothesis, so H_1 is strongly collision-free, which means that the security of H_1 is not less than the security of H_2.

To sum up, the reduced SFRI is equivalent to MD in security. □

***Corollary* 1.1:** SFRI is at least equivalent to MD in security.

Proof:

It is well known from the first part of the proof of theorem 1 that the security of H_2 is not less than the security of H_1, which indicates that the reduced SFRI is at least equivalent to MD in security.

Further, considering the existence of the operator ⇆ in SFRI, we say that SFRI is at least equivalent to MD in security. □

3.1.2 MFRI Structure Being at Least Equivalent to SFRI Structure in Security

We also have the following theorem.

***Theorem* 2:** MFRI is at least equivalent to SFRI in security.

Proof:

For SFRI, $D = f(X_{n+2}, f(X_{n+1}, \leftrightarrows Y_n))$, where $X_{n+1} = X_1$ and $X_{n+2} = X_n$.

For MFRI, $D = f(X_{n+2}, f(X_{n+1}, \leftrightarrows Y_1 \boxplus \dots \boxplus \leftrightarrows Y_n))$, where $X_{n+1} = X_1$ and $X_{n+2} = X_2 \oplus \dots \oplus X_n$.

If we neglect the operations ($\leftrightarrows Y_1 \boxplus \dots \boxplus$) and ($X_2 \oplus \dots \oplus$), then the output of MFRI is reduced to $D = f\,(X_n, f\,(X_1, \leftrightarrows Y_n))$ the same as the output of SFRI, which indicates that the reduced MFRI structure is equivalent to SFRI in security.

Considering the existence of the operations ($\leftrightarrows Y_1 \boxplus \dots \boxplus$) and ($X_2 \oplus \dots \oplus$) in MFRI, we say that MFRI is at least equivalent to SFRI in security. □

3.1.3 It Is Difficult to Find a Message Making Output of MFRI Equal That of SFRI

Further, we have the following property.

***Property* 1:** It is difficult to find a message making the output of MFRI equal that of SFRI.

Proof:

Let the message $M = X_1 \parallel \dots \parallel X_n$.

If we can find in polynomial time a set of values of $X_1, \dots, X_n$ which satisfies the two constraints

$$(\leftrightarrows Y_1 \boxplus \dots \boxplus \leftrightarrows Y_n) = \leftrightarrows Y_n, \text{ and } (X_2 \oplus \dots \oplus X_n) = X_n,$$

the output of MFRI will equal that of SFRI.

According to the definitions of the operators $\boxplus$ and $\oplus$, we see that

$$(\leftrightarrows Y_1 \boxplus \dots \boxplus \leftrightarrows Y_{n-1}) = u2^k, \text{ and } (X_2 \oplus \dots \oplus X_{n-1}) = 0,$$

where $u \in [1, n-1]$ is a positive integer, and k is the bit-length of Y_i.

Let $u = 1$ (if u equals other integers, it does not influence our discussion). In terms of MFRI, the $X_1, \dots, X_{n-1}$ must satisfy the following simultaneous equations:

$$\begin{cases} Y_1 = f(IV, X_1) \\ Y_2 = f(Y_1, X_2) \\ \dots\dots \\ Y_{n-1} = f(Y_{n-2}, X_{n-1}) \\ X_2 \oplus \dots \oplus X_{n-1} = 0 \\ \leftrightarrows Y_1 \boxplus \dots \boxplus \leftrightarrows Y_{n-1} = 2^k \end{cases}$$

Transparently, this equation system contains only two equations substantially, and has $n - 1$ variables $X_1, \dots, X_{n-1}$.

No matter how $X_1, \dots, X_{n-1}$ are sought, the easiest approach must contain the two steps:

① Determine values of any $n - 3$ variables among $X_1, \dots,$ and X_{n-1};

② Compute the values of the other two variables according to the equation system.

Without loss of generality, suppose that the values of $X_1, \ldots, X_{n-3}$ are determined.

Further, the values of X_{n-2} and X_{n-1} need to be sought.

According to $X_2 \oplus \ldots \oplus X_{n-1} = 0$, X_{n-2} can be expressed with the variable X_{n-1}, namely

$$X_{n-2} = X_2 \oplus \ldots \oplus X_{n-3} \oplus X_{n-1}.$$

According to $\leftrightarrows Y_1 \boxplus \ldots \boxplus \leftrightarrows Y_{n-1} = 2^k$, Substitution for $Y_1, \ldots, Y_{n-1}$ yields

$$\leftrightarrows f(IV, X_1) \boxplus \leftrightarrows f(Y_1, X_2) \boxplus \ldots \boxplus \leftrightarrows f(Y_{n-3}, X_{n-2}) \boxplus \leftrightarrows f(Y_{n-2}, X_{n-1}) = 2^k.$$

That is,

$$\leftrightarrows f(IV,X_1) \boxplus \ldots \boxplus \leftrightarrows f(Y_{n-3},X_2 \oplus \ldots \oplus X_{n-3} \oplus X_{n-1}) \boxplus \leftrightarrows f(f(Y_{n-3},X_2 \oplus \ldots \oplus X_{n-3} \oplus X_{n-1}),X_{n-1}) = 2^k.$$

Clearly, seeking X_{n-1} from the above equation is at least equivalent to seeking a preimage of the compression function f. In terms of the one-wayness of f, it is infeasible in polynomial time to seek preimages of f.

Therefore, it is difficult to find a message making the output of MFRI equal that of SFRI. □

Note that property 1 makes us incline to believe MFRI is more secure than SFRI and MD.

3.2 Resistance to the Multi-block Differential Attack

3.2.1 Brief Presentation of the Multi-block Differential Attack

Reference [4], [5], and [6] manifest the multi-block near differential attack on the hash functions MD4, MD5, SHA-0, and SHA-1. This attack consists of the three steps:

① Find out a set of collision differential characteristics for M and M' which are expected to produce a collision.

② Derive a set of sufficient conditions which are described by the bits of the chaining variables, and ensure that the collision differential characteristics hold.

③ Modify the random message M through the single-step / multi-step or single-message / multi -message method in order to make almost all the sufficient conditions be satisfied.

Assume M is partitioned into n m-bit blocks $X_1, \ldots, X_n$, and the iterative outputs are $Y_1, \ldots, Y_n = D$.

Assume M' is partitioned into n m-bit blocks $X'_1, \ldots, X'_n$, and the iterative outputs are $Y'_1, \ldots, Y'_n = D'$, where $n \geq 2$.

According to [5] and [6], define the message differential as $\Delta X_i = X'_i - X_i$, and the iterative output differential as $\Delta Y_i = Y'_i - Y_i$.

Note that a differential is computed by modular integer subtraction '–' in [5] and [6] while it is computed by exclusive or '⊕' in other references. Obviously, the combination of these two sorts of differentials can bring more information to attackers.

For MD, assume collision differential characteristics are $\Delta X_1, \ldots, \Delta X_n$, and $0 = \Delta Y_0, \Delta Y_1, \ldots, \Delta Y_n = 0$.

It is should be noted that because the same compression function f is used when two different messages are hashed, the initial values of iteration are the same, namely $\Delta Y_0 = 0$. The $\Delta Y_n = 0$ indicates that the collision (M, M') is found out, and it is a goal which the attackers try to achieve.

ΔY_i is also the chaining variable difference. In terms of a concrete compression function, the attackers may set more detailed step-chaining variable differentials and round- chaining variable differentials [5][6].

3.2.2 MFRI Leading Block Modification to a Joint Event

In MFRI, let the collision differential characteristics be $\Delta X_1, \ldots, \Delta X_n$, $0 = \Delta Y_0, \Delta Y_1, \ldots, \Delta Y_n$, and the input chaining variable of the $(n + 1)$-th iteration be Y_n^{d}, then $Y_n^{\mathrm{d}} = \leftrightarrows Y_1 \boxplus \ldots \boxplus \leftrightarrows Y_n$, where the superscript ‘d’ signifies time delay. Therefore,

$$\Delta Y_n^{\mathrm{d}} = (\leftrightarrows Y'_1 \boxplus \ldots \boxplus \leftrightarrows Y'_n) - (\leftrightarrows Y_1 \boxplus \ldots \boxplus \leftrightarrows Y_n)$$
$$= \Delta\leftrightarrows Y_1 \boxplus \ldots \boxplus \Delta\leftrightarrows Y_n \neq \leftrightarrows\Delta Y_1 \boxplus \ldots \boxplus \leftrightarrows\Delta Y_n.$$

For example, when $A = 11010100$, $A' = 00101011$, and $\Delta A = 10101001$, there are $\leftrightarrows A = 00101011$, $\leftrightarrows A' = 11010100$, and $\Delta\leftrightarrows A = 01010111$, and so $\Delta\leftrightarrows A \neq \leftrightarrows\Delta A$.

This brings extra difficulties to the attackers who employ the differential analysis method.

From [5] and [6], it is not difficult to understand that if there are not the multiple feedbacks, the modification to every block X_i is an independent event, and when the modification is made, it is feasible to consider ΔX_i and ΔY_i only relevant to the block X_i but not to other blocks. However, when the multiple feedbacks exist, due to

$$\Delta Y_n^{\mathrm{d}} = (\leftrightarrows Y'_1 \boxplus \ldots \boxplus \leftrightarrows Y'_n) - (\leftrightarrows Y_1 \boxplus \ldots \boxplus \leftrightarrows Y_n),$$

the modification to every block X_i will influence the corresponding Y_i, and further do ΔY_n^{d}. Thereby, the modification to every block X_i changes into an joint event from an individual independent event.

Assume that through message modification techniques the attackers can decrease the time complexity of a block near collision $O(2^{k_i})$. In terms of [5] and [6], in MD, the modification to every block is an independent event, and hence, the complexity of producing the n-block message collision is $O(2^{k_1} + \ldots + 2^{k_n})$. However, in MFRI, the modification to every block become a part of the joint event, and hence, the probability that two n-block messages produce a collision is $1 / (2^{k_1} \ldots 2^{k_n})$, namely, the complexity of producing the message collision increases to $O(2^{k_1 + \ldots + k_n})$.

3.2.3 MFRI Leading Message Modification to an Endless Loop

To ensure that the differential characteristics being set holds, every block has a set of sufficient conditions derived from f and ΔY_i, where ΔY_i is a chaining variable differential. For the hash functions MD4 and MD5, the length of every chaining variable is 128 bits, is exactly one of four 32-bit words. Therefore, in fact, every chaining variable consists of the four word variables a, b, c, and d. Because every block-iteration consists of several round-iterations, and every round-iteration consists of several step-iterations, the variables a, b, c, d may be further divided into $a_1, b_1, c_1, d_1, \ldots, a_s, b_s, c_s, d_s$ in every block-iteration process. For example, in MD4, $s = 12$, and in MD5, $s = 16$. The values of a_i, b_i, c_i, d_i in the sufficient conditions are expressed with 0, 1 or those prior to a_i, b_i, c_i, d_i. For SHA-1, its chaining variable is composed of a, b, c, d, and e five word variables. The values of the sufficient conditions for a block collision are expressed with 0, 1 or those prior to a_i.

In MFRI, because of feedback and $X_{n+1} = X_1$, the second modification to X_1 is needed. However, the second modification will surely influence the result of the first modification, that is, break the sufficient conditions satisfied and change the value of the chaining variable Y_1, and further, cause Y_2, …, Y_n and the feedback Y_n^{d} to produce alteration. Thus, the attackers need to modify X_1 once more. In this way, MFRI will lead the modification to X_1 and other blocks to an endless loop.

3.2.4 MFRI Leading Sufficient Conditions for Collision to Incompatibility

Due to $X_{n+1} = X_1$, for MFRI, a key problem is whether the two sets of sufficient conditions described with the chaining variables respectively in the 1-st iteration and (n + 1)-th iteration are compatible or not. If they are compatible, the two sets of sufficient conditions can be deduced theoretically. If they are contrary to each other, the two sets of sufficient conditions is radically impossibly deduced, and thus it is not ensured that the differential characteristics will hold and that the collision for two messages can be found.

Suppose that $\mathit{\Lambda} = \{a_{1,1}, \ldots, a_{1,32}\}$ is the set of 32 bits of the variable a_1, then the size of its power set is $|Ƥ(\mathit{\Lambda})| = 2^{32}$. Let $P_{\cap\mathrm{NE}}$ denote the probability that the intersection of any two nonempty subsets of $\mathit{\Lambda}$ is empty, then

$$\begin{aligned} P_{\cap\mathrm{NE}} &= (\mathrm{C}_{32}^{1}(2^{31}-1) + \ldots + \mathrm{C}_{32}^{31}(2^{1}-1)) / (2\mathrm{C}_{|Ƥ(\mathit{\Lambda})|}^{2}) \\ &= (\mathrm{C}_{32}^{1}2^{31} + \ldots + \mathrm{C}_{32}^{31}2^{1} - \mathrm{C}_{32}^{1} - \ldots - \mathrm{C}_{32}^{31}) / (2^{32}(2^{32}-1)) \\ &= (3^{32} - 2^{32} - 1 - (2^{32}-2)) / (2^{32}(2^{32}-1)) \\ &= (1.6458 \times 2^{50}) / (2^{32}(2^{32}-1)) < (2^{51}) / (2^{32}(2^{32}-1)) \approx 1/2^{13}. \end{aligned}$$

Thereby, the probability that the intersection of any two nonempty subsets of $\mathit{\Lambda}$ is nonempty is greater than $(1 - 1/2^{13})$, which means that probability at least 1 bit of the condition variable a_1 produces overlap in the 1-st iteration and (n + 1)-th iteration is greater than $(1 - 1/2^{13})$.

We may suppose that $a_{1,1}$ overlaps, let $a_{1,1}^{1}$ denote the condition value of $a_{1,1}$ in the 1-th iteration, and let $a_{1,1}^{n+1}$ denote the condition value of $a_{1,1}$ in the (n + 1)-th iteration. If $a_{1,1}^{1} = a_{1,1}^{n+1} = 0$ or 1, it indicates the two sets of sufficient conditions are compatible; otherwise the two sets of sufficient conditions are incompatible, that is, such two sets of sufficient conditions may impossibly exist simultaneously. Obviously, if the intersection contains only 1 bit, the probability of being incompatible is 1/2. If the intersection contains 2 bits, the probability of being incompatible is (1 – 1/4). Suppose that P_{a_1}, P_{b_1}, P_{c_1}, P_{d_1} = 1 / 2 or 1 represent respectively the probabilities that the intermediate chaining variables a_1, b_1, c_1, d_1 are condition-compatible in the 1-th iteration and (n + 1)-th iteration. Then, for a_1, b_1, c_1, and d_1, the probability that the conditions are incompatible is $(1 - 1 / (P_{a_1}P_{b_1}P_{c_1}P_{d_1}))$. For the other intermediate chaining variables a_2, b_2, c_2, d_2 …, there exist similar conclusions.

The above analysis manifests that in two different iterations of the same block, the probability that the condition bits produce overlap is close to 1, and the probability that the values of the overlapping bits are incompatible is greater than 1/2.

3.3 Ineffectiveness of the D-way Second Preimage Attack

Joux puts forward a attack method called D-way which is employed for seeking the second preimage of an output of a hash function based on MD in [8]. For a given hash target value $Y = H(M) \in \{0, 1\}^k$, the attackers first find 2^d collisions on d-block messages $M_1, \ldots, M_{2^d}$ making $H_d = H(M_1) = \ldots = H(M_{2^d})$. Then, find the block X_{d+1} such that $f(H_d, X_{d+1}) = Y$. In this way, the attackers succeed in seeking the second preimage with the message M. In terms of [8], the time complexity of this attack is $O(d2^{k/2} + 2^k)$.

For a hash function based on MFRI, because there are $X_{d+1} = X_1$, $X_{d+2} = X_2 \oplus \ldots \oplus X_d$, and $X'_{d+1} = X'_1$, $X'_{d+2} = X'_2 \oplus \ldots \oplus X'_d$, even though $\Delta Y_1 = \ldots = \Delta Y_d = 0$, it can not be ensured that $\Delta Y_{d+1} = 0$ and $\Delta Y_{d+2} = 0$. That is, it is intractable to find out two d-block messages M_1 and M_2 for collision by the birthday attack [8]. Therefore, the D-way method is ineffective on hash functions based on MFRI.

4 Performance Analysis of the Hash Ring-iterative Structure

For the same message M, MFRI is two f mapping operations, n reverse operations, n modular addition operations and $(n - 1)$ exclusive OR operations more than MD. Reverse, modular addition and exclusive OR are fundamental operations, and they can not expend too much time. Hence, MFRI has comparatively fast operation speed. The two extra variables in memory space need to be increased respectively for the feedforward and feedback values. The initial values of these two variables may be set to zero. Then, the feedforward variable admit $X_2, \ldots$, and X_n one by one by exclusive OR, and the feedback variable admit $\leftrightarrows Y_1, \ldots$, and $\leftrightarrows Y_n$ one by one by modular addition.

At present, the tree structure and zipper structure for hash functions are also believed to be more secure than MD [9][10]. However, MFRI is more applicable than the tree structure since the compression mapping f in any existing hash function may be transplanted into MFRI with no change, and is more efficient than the zipper structure since the number of time of operation on the mapping f in MFRI is roughly half as many as in the zipper structure.

5 Conclusions

At the time every iteration output is fed back, first to do a reversal transform is important, which makes it impossible that ΔY_n^{d} is derived directly from $\Delta Y_1, \ldots$, and ΔY_n.

It is known from section 3.1 that there is some comparability between MFRI and MD. The last two extra blocks in MFRI may be regarded as an extension of the MD padding. Therefore, for the same compression f, if the hash output of MD is uniform, independent and random, the hash output of MFRI is also uniform, independent and random. If MD can cause the avalanche effect of the hash output, MFRI can also cause the avalanche effect of the hash output. Note that for the input message *X*, only if it contains at least two blocks does MFRI take effect.

For the existing hash functions ─ MD5 and SHA-1 for example, if their compression functions are extracted and transplanted into MFRI, the preceding analysis shows that the existing attack methods will be ineffective on the newly forming hash functions.

Acknowledgment

The authors would like to thank Zhaozhi Zhang, and Xinxin Niu for their important counsel, suggestions, and corrections.

References

[1] A. J. Menezes, P. Van Oorschot, and S. Vanstone, *Handbook of Applied Cryptography*, CRC Press, London, UK, ch. 9, 1997.

[2] Ivan Damgard, "A design principle for hash functions", *Advances in Cryptology: CRYPTO 89*, Springer-Verlag, Berlin, DE, pp. 416-427, 1989.

[3] Ralph Merkle, "One way hash functions and DES", *Advances in Cryptology: CRYPTO 89*, Springer-Verlag, Berlin, DE, pp. 428-446, 1989.

[4] Eli Biham, Rafi Chen, Antoine Joux etc, "Collisions of SHA-0 and Reduced SHA-1", *Advances in Cryptology - EUROCRYPT 05*, Springer-Verlag, Berlin, DE, pp. 36-57, 2005.

[5] Xiaoyun Wang, Yiqun Lisa Yin, and Hongbo Yu, "Finding collisions in the full SHA-1", *Advances in Cryptology - CRYPTO 05*, Springer-Verlag, Berlin, DE, pp. 17-36, 2005.

[6] Xiaoyun Wang, Hongbo Yu, "How to Break MD5 and Other Hash Functions", *Advances in Cryptology - EUROCRYPT 05*, Springer-Verlag, Berlin, DE, pp. 19-35, 2005.

[7] Praveen Gauravaram, William Millan, Ed Dawson etc, "Constructing Secure Hash Functions by Enhancing Merkle-Damgård Construction", *Conf. on ISP 06*, LNCS v4058, Springer-Verlag, Berlin, DE, pp. 407-420, 2006.

[8] Antoine Joux, "Multicollisions in Iterated Hash Functions. Application to Cascaded Constructions", *Advances in Cryptology-CRYPTO 04*, Springer-Verlag, Berlin, DE, pp. 306-316, 2004.

[9] Oded Goldreich, *Foundations of Cryptography Volume II, Basic Applications*, Cambridge University Press, Cambridge, UK, pp. 521-523, 2004.

[10] Moses Liskov, *Constructing Secure Hash Functions from Weak Compression Functions: The Case for Non-Streamable Hash Functions*, available: www.cs.wm.edu/~mliskov/hash.pdf, 2006.